Acta Crystallographica Section F
**Structural Biology and Crystallization Communications**
ISSN 1744-3091

# Overproduction, crystallization and preliminary X-ray characterization of Abn2, an endo-1,5-α-arabinanase from *Bacillus subtilis*

**Daniele de Sanctis,[a]* Isabel Bento,[a] José Manuel Inácio,[a] Sónia Custódio,[a] Isabel de Sá-Nogueira[a,b] and Maria Arménia Carrondo[a]**

[a]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Avenida da República–EAN, 2780-157 Oeiras, Portugal, and [b]Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, 2829-516 Caparica, Portugal

Correspondence e-mail: desanctisd@gmail.com

Received 17 April 2008
Accepted 29 May 2008



Two *Bacillus subtilis* extracellular endo-1,5-α-L-arabinanases, AbnA and Abn2, belonging to glycoside hydrolase family 43 have been identified. The recently characterized Abn2 protein hydrolyzes arabinan and has low identity to other reported 1,5-α-L-arabinanases. Abn2 and its selenomethionine (SeMet) derivative have been purified and crystallized. Crystals appeared in two different space groups: $P1$, with unit-cell parameters $a$ = 51.9, $b$ = 57.6, $c$ = 86.2 Å, $\alpha$ = 82.3, $\beta$ = 87.9, $\gamma$ = 63.6°, and $P2_12_12_1$, with unit-cell parameters $a$ = 57.9, $b$ = 163.3, $c$ = 202.0 Å. X-ray data have been collected for the native and the SeMet derivative to 1.9 and 2.7 Å resolution, respectively. An initial model of Abn2 is being built in the SeMet-phased map.

## 1. Introduction

Arabinan is an L-arabinose homoglycan present in plant tissues that is generally associated with pectins. This polysaccharide is composed of α-1,5-linked L-arabinofuranosyl units, some of which are substituted with α-1,3- and α-1,2-linked chains of L-arabinofuranosyl residues (Beldman *et al.*, 1997). The two major enzyme families that hydrolyse arabinan are α-L-arabinofuranosidases (EC 3.2.1.55; AFs) and endo-1,5-α-L-arabinanases (EC 3.2.1.99; ABNs). AFs remove arabinose side chains, allowing ABNs to attack the glycosidic bonds of the arabinan backbone, releasing a mixture of arabino-oligosaccharides and L-arabinose as the products of the reaction. Arabinanases have several applications, in particular in the food industry, such as food technology and nutritional medical research (Beldman *et al.*, 1997).

*Bacillus subtilis*, a Gram-positive endospore-forming bacterium which participates in plant-biomass degradation, synthesizes two extracellular ABNs belonging to glycoside hydrolase family 43 (GH43; Coutinho & Henrissat, 1999), AbnA and Abn2 (YxiA), which hydrolyze sugar-beet arabinan (branched), linear α-1,5-L-arabinan and pectin (Inácio & de Sá-Nogueira, 2008; Leal & de Sá-Nogueira, 2004). The three-dimensional structures of three different 1,5-α-L-arabinanases, Arb43A from *Cellvibro japonicus* (Nurizzo *et al.*, 2002), *Bs*Arb43A (AbnA) from *B. subtilis* (Proctor *et al.*, 2005) and ABN-TS from *B. thermodenitrificans* (Yamaguchi *et al.*, 2005), have been determined by X-ray crystallography. These enzymes, which vary in length from 323 to 347 amino acids, display a unique motif consisting of a five-bladed β-propeller fold. Since the recently characterized Abn2 from *B. subtilis* (Inácio & de Sá-Nogueira, 2008) shows less than 23% amino-acid identity to these 1,5-α-L-arabinanases and is a much larger enzyme (469 residues) than other related arabinases, the determination of its three-dimensional structure should provide valuable information concerning the structural features of glycoside hydrolase family 43 arabinanases. Here, we report the crystallization and preliminary X-ray analysis of the GH family 43 Abn2 from *B. subtilis*.

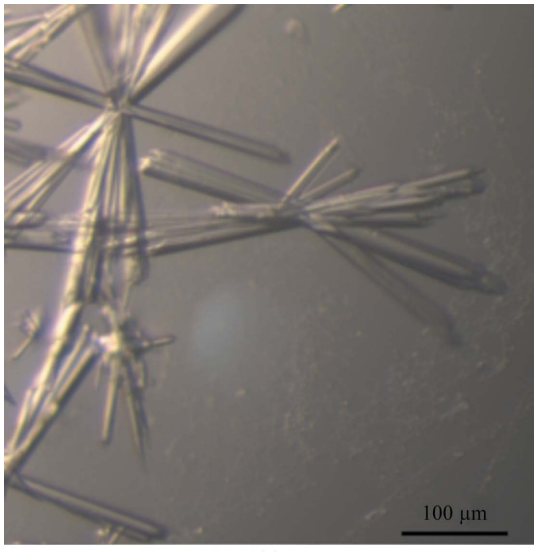
## 2. Experimental procedures

### 2.1. Overproduction and purification

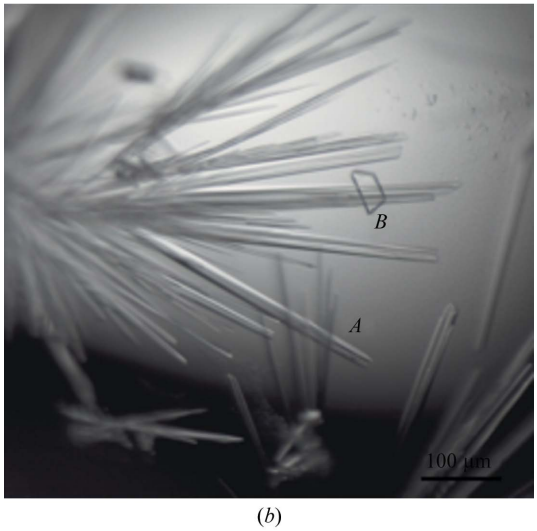
The overproduction and purification of Abn2 in *Escherichia coli* BL21 (DE3) pLysS has been described previously (Inácio & de Sá-Nogueira, 2008). Briefly, the *abn2* allele (GenBank accession No.

EU373814) was amplified by PCR with primers ARA237 (5′-GGC-GAATTGTT**CATATG**TTCAACCG-3′) and ARA238 (5′-CGCTT-CTCC**CTCGAG**TTTAGATCCC-3′) using chromosomal DNA of wild-type strain *B. subtilis* 168 T$^+$ as template. The resulting 1568 bp DNA fragment was digested with *Nde*I–*Xho*I unique restriction sites introduced by the primers (bold sequence) and cloned into the same sites of pET30a(+) (Novagen). In the resulting plasmid pZI39, the wild-type primary sequence of the Abn2 is fused directly to the carboxyl-terminal sequence (LEHHHHHH) encoded by pET30a(+).

*E. coli* BL21 (DE3) pLysS cells harbouring pZI39 were grown at 310 K and 160 rev min$^{-1}$ in 1 l LB with an appropriate antibiotic selection. When the $OD_{600}$ reached 0.6, expression of Abn2 was induced by the addition of 1 m*M* IPTG. The culture was grown for an additional 4 h at 310 K and 160 rev min$^{-1}$. Cells were harvested by centrifugation at 277 K and 8000*g* for 10 min. All subsequent steps were carried out at 277 K. The periplasmic protein fraction (PPF) was prepared by osmotic shock and loaded onto a 1 ml HisTrap column (Amersham Phamarcia Biotech). The Ni-bound proteins were eluted with a discontinuous imidazole gradient and fractions containing Abn2 that were more than 95% pure were dialyzed overnight against a buffer containing 100 m*M* Tris pH 8.0, 100 m*M* NaCl, 10%(*v*/*v*) glycerol. Purified protein was then concentrated using an Amicon Ultra 15 Centrifugal filter (Millipore) to a final concentration of 10.9 mg ml$^{-1}$ and stored in the same buffer at 193 K. Seleno-L-methionine-labelled protein was overexpressed in an auxotrophic *E. coli* strain B834 (DE3) harbouring pZI39. A colony was used to inoculate 10 ml LB with an appropriate antibiotic selection and the culture was incubated for 6 h at 150 rev min$^{-1}$ and 310 K. The culture was then diluted in a 1:20 ratio with 100 ml of SeMet Base media (Molecular Dimensions Ltd), L-methionine, Nutrient Mix (Molecular Dimensions Ltd) and an appropriate antibiotic selection. After overnight incubation the cell pellet was separated by centrifugation at 3000*g* and 298 K for 10 min, washed to remove residual L-methionine and resuspended in 1 l of the same medium containing L-seleno-methionine (SeMet; Molecular Dimensions Ltd) instead of L-methionine. When the $OD_{600}$ reached 0.75, induction was performed by the addition of 1 m*M* IPTG. The culture was grown for an additional 3 h and cells were then harvested. Preparation of the PPF and purification of SeMet-Abn2 was carried out as described above for the native protein. Purified protein was concentrated to a final concentration of 15.4 mg ml$^{-1}$ using an Amicon Ultra 15 Centrifugal filter (Millipore) and stored in the same buffer as the native protein at 193 K.




(*a*)




(*b*)

**Figure 1**
(*a*) SeMet-derivative crystals of Abn2. The dimensions of a typical crystal fragment used in data collection were 0.05 × 0.05 × 0.2 mm. (*b*) Crystals of the native Abn2 crystals. Two crystalline forms are present. Crystal form *A* belongs to space group *P*1 and has typical dimensions of 0.02 × 0.06 × 0.1 mm; crystal form *B* belongs to space group $P2_12_12_1$.

**Table 1**
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

| | SeMet | Wild type |
|---|---|---|
| Wavelength (Å) | 0.97865 | 1.03320 |
| Space group | $P2_12_12_1$ | *P*1 |
| Unit-cell parameters (Å, °) | *a* = 57.9, *b* = 163.3, *c* = 202.0 | *a* = 51.9, *b* = 57.6, *c* = 86.2, $\alpha$ = 82.3, $\beta$ = 87.9, $\gamma$ = 63.6 |
| Resolution limits (Å) | 127.0–2.70 (2.85–2.70) | 85.4–1.90 (2.00–1.90) |
| $R_{merge}$† | 0.093 (0.190) | 0.047 (0.136) |
| $R_{p.i.m.}$‡ | 0.056 (0.121) | 0.045 (0.036) |
| Total No. of reflections | 365808 (46766) | 134405 (17357) |
| No. of unique reflections | 53601 (7661) | 67116 (8980) |
| Mean $I/\sigma(I)$ | 17.6 (7.4) | 14.5 (6.4) |
| Completeness (%) | 99.9 (99.6) | 95.7 (87.1) |
| Multiplicity | 6.8 (6.1) | 2.0 (1.9) |
| $R_{ano}$§ | 0.064 (0.112) | |
| Anomalous completeness (%) | 99.1 (95.0) | |
| Anomalous multiplicity | 3.5 (3.2) | |

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{p.i.m.} = \sum_{hkl}[N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. § $R_{ano} = \sum_h |\langle I(h)\rangle - \langle I(-h)\rangle| / \sum_h [\langle I(h)\rangle + \langle I(-h)\rangle]$, where $\langle I(h)\rangle$ indicates the mean intensity of symmetry-related reflections of the same Bijvoet mates *h*.

## 2.2. Crystallization

Preliminary crystallization screens were carried out with the native protein using a sitting-drop vapour-diffusion experimental setup (100 nl protein solution added to 100 nl reservoir solution equilibrated against 100 µl reservoir solution) using a Cartesian Nanodrop Robot to set up 96-well plates at 294 K. Crystals were found in condition No. 22 [0.1 *M* HEPES pH 7.5, 70%(*v*/*v*) MPD] of Classics Screen (Nextal). Crystal growth was scaled up by hanging-drop vapour diffusion at 294 K in 48-well plates (1 µl protein solution and 1 µl reservoir solution equilibrated against 200 µl reservoir solution), obtaining the best shaped crystals in 65%(*v*/*v*) MPD, 0.1 *M* Tris pH 8.5. The crystals that appeared showed two distinct morphologies (Fig. 1). Final crystal dimensions were 40 × 50 × 200 µm for crystal form *A* and 20 × 60 × 100 µm for crystal form *B*. The SeMet-protein

crystallization experiment was set up using the same conditions as used for the native protein, but crystals appeared only in one crystalline form (Fig. 1). The best-shaped crystals were obtained with 62%(*v*/*v*) MPD and 0.1 *M* Tris pH 8.4.

### 2.3. X-ray analysis

Crystals were transferred into a stabilizing solution containing 70%(*v*/*v*) MPD and 0.1 *M* Tris pH 8.5 and subsequently flash-cooled in liquid nitrogen. Native protein crystals with different morphologies belonged to different space groups: crystal form *A* belonged to space group $P2_12_12_1$, with unit-cell parameters $a$ = 57.81, $b$ = 164.48, $c$ = 203.31 Å, and can accommodate four molecules in the asymmetric unit with a Matthews coefficient $V_M$ = 2.29 Å$^3$ Da$^{-1}$ and a solvent content of 46%, while crystal form *B* belonged to space group *P*1, with unit-cell parameters $a$ = 51.97, $b$ = 57.64, $c$ = 86.27 Å, $\alpha$ = 82.31, $\beta$ = 87.97, $\gamma$ = 63.66°, which is compatible with the presence of two molecules in the asymmetric unit, corresponding to a Matthews coefficient $V_M$ = 2.17 Å$^3$ Da$^{-1}$ and a solvent content of 43% (Matthews, 1985). As the *P*1 crystals were the better diffracting crystal form, a data set was collected from these native crystals to 1.9 Å resolution at ID29, ESRF (Grenoble) at 100 K in 180 non-overlapping 1° oscillation images with a crystal-to-detector distance of 265.56 mm. Data were indexed with *MOSFLM* (Leslie, 2006) and scaled with *SCALA* (Evans, 2006).

The SeMet derivative of Abn2 only crystallized in one crystal form, which was similar to form *A* of the wild-type protein. SeMet-derivative crystals diffracted to 2.7 Å resolution and belonged to space group $P2_12_12_1$, with unit-cell parameters $a$ = 57.84, $b$ = 163.22, $c$ = 201.85 Å, which can accommodate four molecules (each containing 13 selenomethionine residues) in the asymmetric unit, corresponding to a Matthews coefficient of 2.26 Å$^3$ Da$^{-1}$ and a solvent content of 46%.

SAD data were collected from an SeMet-labelled crystal at 100 K at the BM14 Station at ESRF at a wavelength of 0.97865 Å, which was chosen after an absorption scan measurement to maximize the anomalous signal of Se atoms. Oscillation steps of 0.2° over a range of 180° were applied with a crystal-to-detector distance of 291.89 mm. SeMet-derivative crystal diffraction images were indexed with *MOSFLM* (Leslie, 2006) and scaled with *SCALA* (Evans, 2006).

### 2.4. Structure solution

The program *SOLVE* (Terwilliger & Berendzen, 1999) was used to locate 31 of the 42 Se atoms in the asymmetric unit and to perform initial phasing to a figure of merit of 33% ($Z$ score 54.3). Subsequently, *RESOLVE* (Terwilliger, 2000) was used to perform statistical density modification with noncrystallographic symmetry averaging to a final figure of merit of 72%. In the final map, autotracing using *RESOLVE* (Terwilliger, 2003) succeeded in building some main features of the four independent molecules in the asymmetric unit, resulting in the placement of 765 residues out of 1876. Manual model building with *Coot* (Emsley & Cowtan, 2004) is in progress.

This work was partially supported by grant No. POCI/AGR/60236/2004 from Fundação para a Ciência e Tecnologia (FCT) and FEDER to IdS-N and fellowship SFRH/BD/18238/2004 from FCT to JMI.

## References

Beldman, G., Schols, H. A., Pitson, S. M., Searle-van Leeuwen, M. J. F. & Voragen, A. G. J. (1997). *Advances in Macromolecular Carbohydrate Research*, Vol. 1, edited by R. J. Sturgeon, pp. 1–64. Stamford: JAI Press.
Coutinho, P. M. & Henrissat, B. (1999). *Recent Advances in Carbohydrate Bioengineering*, edited by H. J. Gilbert, G. J. Davies, B. Henrissat & B. Svensson, pp. 3–12. Cambridge: The Royal Society of Chemistry.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Inácio, J. M. & de Sá-Nogueira, I. (2008). *J. Bacteriol.* **190**, 4272–4280.
Leal, T. F. & de Sá-Nogueira, I. (2004). *FEMS Microbiol. Lett.* **241**, 41–48.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Matthews, B. W. (1985). *Methods Enzymol.* **114**, 176–187.
Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., Taylor, E. J., Gilbert, H. J. & Davies, G. J. (2002). *Nature Struct. Biol.* **9**, 665–668.
Proctor, M. R., Taylor, E. J., Nurizzo, D., Turkenburg, J. P., Lloyd, R. M., Vardakou, M., Davies, G. J. & Gilbert, H. J. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 2697–2702.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 38–44.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Yamaguchi, A., Tada, T., Wada, K., Nakaniwa, T., Kitatani, T., Sogabe, Y., Takao, M., Sakai, T. & Nishimura, K. (2005). *J. Biochem.* **137**, 587–592.